(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 874 048 A2

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
28.10.1998 Bulletin 1998/44

(21) Application number: **98303187.3**

(22) Date of filing: **24.04.1998**

(51) Int Cl.[6]: **C12N 15/12**, C12N 5/20, C12P 21/08, C07K 14/47, C07K 16/18

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **25.04.1997 JP 121578/97**
14.04.1998 JP 117873/98

(71) Applicant: **KABUSHIKI KAISHA HAYASHIBARA SEIBUTSU KAGAKU KENKYUJO**
Okayama-shi Okayama (JP)

(72) Inventors:
- **Ariyasu, Toshio**
  Okayama-shi, Okayama (JP)
- **Nakamura, Shuji**
  Okayama-shi, Okayama (JP)
- **Orita, Kunzo**
  Okayama-shi, Okayama (JP)

(74) Representative: **Daniels, Jeffrey Nicholas et al**
Page White & Farrer
54 Doughty Street
London WC1N 2LS (GB)

(54) Human hedgehog protein

(57) Disclosed are a novel hedgehog protein, i.e., a Desert hedgehog protein of human origin including mature and precursor forms, a DNA encoding the protein, a monoclonal antibody recognizing the protein, a process for producing the protein, and a method for detecting the protein. The hedgehog protein is useful in establishment of hybridomas which produce antibodies recognizing the protein, and the monoclonal antibody is useful in detection and purification of the protein. The hedgehog protein, DNA, and monoclonal antibody of this invention have efficacy in elucidation of hereditary morphological abnormalities in humans to establish their treatments and diagnoses.

EP 0 874 048 A2

Printed by Jouve. 75001 PARIS (FR)

protein was electrophoresed. Numbers on left side of each lane mean the molecular weights of molecular weight markers in a unit of kilodaltons and indicate their positions after electrophoresis.

In Figure 3, closed circles represent the results of detecting human Desert hedgehog protein, and closed squares represent the results of detecting human Sonic hedgehog protein.

This invention relates to a novel hedgehog protein, more particularly, a Desert hedgehog protein of human origin. The present hedgehog protein contains a part or the whole of the amino acid sequence of SEQ ID NO:1, which may bear a homology, usually about 80%, to mouse Desert hedgehog protein at amino acid sequence level. Examples of the present hedgehog protein are a mature form of human Desert hedgehog protein that contains the amino acid sequence of SEQ ID NO:1 and a precursor form of human Desert hedgehog protein that has the amino acid sequence of SEQ ID NO:2 or 3, which contains the amino acid sequence of SEQ ID NO:1. The present hedgehog protein further includes other types of proteins with amino acid sequences as illustrated above where one or more amino acids are deleted or replaced with other ones, one or more amino acids are added, or saccharide chains are linked, so far as they contain the amino acid sequence as mentioned above. The present hedgehog protein shall not be restricted to those obtained from specific sources and by specific preparation methods, threrefor it include natural proteins obtained from cultures of established cell lines, recombinant proteins obtained by recombinant DNA techniques, and synthetic polypeptides obtained by way of peptide synthesis.

The DNA of this invention includes any DNAs which encode such hedgehog protein, regardless of their sources or origins. Thus the DNA of this invention include those from natural sources as well as those artificially modified or chemically synthesized, as far as they encode the hedgehog protein of this invention. Generally in this field, in case of artificially expressing DNAs which encode proteins, one may replace one or more nucleotides in the DNAs with different nucleotides and/or link appropriate nucleotide sequences thereto with the purpose of improving their expression efficiency and/or the physiological and physicochemical properties of the protein. Such modification are feasible in the DNA of this invention. More particularly, one can link, for example, to the 5'- and/or 3'-termini of the DNA as described above, recognition sites for appropriate restriction enzymes, initiation codons, termination codons, promotors and/or enhancers, as far as the final protein products do retain prescribed properties. Thus, the wording "DNA" as referred to in this invention shall mean, in addition to those which encode the above-mentioned proteins, those which are complementary thereto, and those where one or more nucleotides have been replaced with different nucleotides while conserving the encoding amino acid sequence.

Such a DNA can be obtained from natural by screening of human cells, for example, mammalian cells including epithelial cells, endothelial cells, interstitial cells, chondrocytes, monocytes, granulocytes, lymphocytes, neurocytes, and established cell lines from them of human origin, based on a hybridization with a DNA as a probe which encodes at least a part of the amino acid sequence of human Desert hedgehog protein of this invention, for example, the amino acid sequence of SEQ ID NO:1. Such screening can be achieved with conventional methods commonly used in this field such as PCR, RT-PCR, screening cDNA libraries, screening genomic libraries and/or modified methods thereof. Examples of preferred cells are established cell lines including ARH-77 cell, ATCC CRL-1621, K-562 cell, ATCC CCL-243, and KU-812 cell, an cell line reported by K. Kishi, in *Leukemia Research,* Vol.9, pp.381-390 (1985), and bone mallow cells. The DNA of this invention thus obtained usually contains a part or the whole of the nucleotide sequence of SEQ ID NO:4. For example, from ARH-77 cell, ATCC CRL-1621, a DNA encoding a mature form of human Desert hedgehog protein that contains the nucleotide sequence of SEQ ID NO:4 or a DNA encoding a precursor form of human Desert hedgehog protein that has the nucleotide sequence of SEQ ID NO:5 or 6, which contains the nucleotide sequence of SEQ ID NO:4, can be obtained. The present DNA can also be obtained by conventional chemical synthesis. The DNA of this invention, once obtained in any manner, can be easily amplified to desired level by methods of PCR or those using autonomously replicable vectors.

The DNA of this invention includes those in the forms of recombinant DNAs where the DNA, encoding the present hedgehog protein, is inserted into autonomously replicable vectors. The recombinant DNAs can be relatively-easily obtained by using conventional recombinant DNA techniques, once the desired DNA is obtained. Examples of the vectors feasible in this invention are plasmid vectors including pGEX-2T, pGEX-4T-1, pKK223-3, pcDNAI/Amp, BCM-GSNeo, pcDL-SRa, pKY4, pCDM8, pCEV4, and pME18S. The autonomously replicable vectors usually comprise nucleotide sequences suitable for the DNA expression in respective hosts, for example, promotors, enhancers, replication origins, terminators for transcription, splicing sequences, and/or sequences for selection markers. As the promotor, using a heat shock protein promotor or the interferon-a promotor disclosed in Japanese Patent Kokai No.163,368/95 by the same applicant makes it possible to regulate the present DNA expression in the transformants by external stimuli.

To insert the DNA of this invention, conventional methods comonly used in this field can be used. More particularly, a gene containing the DNA of this invention and an autonomously replicable vector are first cleaved with restriction enzymes and/or ultrasonication, then the resulting DNA and vector fragments are ligated. Ligation of the DNA and vector fragments become much easier when the genes and vectors are digested with restriction enzymes specific to particular nucleotides, for example, *Acc*I, *Bam*HI, *Bst*XI, *Eco*RI, *Hind*III, *Not*I, *Pst*I, *Sac*I, *Sal*I, *Sma*I, *Spe*I, *Xba*I and *Xho*I. To ligate the DNA and vector fragments, they can be first annealed, if necessary, and then exposed to DNA ligase

transformants and culture conditions, is usually one microgram to 100 mg per liter.

Furthermore, in the process for preparing the hedgehog protein of this invention, the DNA expression step can include a step of culturing cells which express the hedgehog protein, for example, established human cell lines ARH-77 cell, ATCC CRL-1621, K-562 cell, ATCC CCL-243, and KU-812 cell, described by K. Kishi et al., in *Leukemia Research,* Vol.9, pp.381-390 (1985). By culturing such cells in culture media suitable for respective cells, for example, 199 medium, DMEM medium, Ham's F12 medium, IMDM medium, MCDB104 medium, MCDB153 medium, MEM medium, RD medium, RITC80-7 medium, RPMI-1630 medium, RPMI-1640 medium, and WAJC404 medium similarly as in culturing of the trasnformants using animal host cells as mentioned above, then the culture containing the present hedgehog protein can be obtained. The content of the present protein in the cultures, which may differ depending on the types of the cells and culture conditions, is usually one nanogram to one milligram per liter.

The culture products obtained in these manners can be first subjected to ultrasonication, cell-lytic enzyme and/or detergent to disrupt cells, if necessary, the present hedgehog protein can be separated from the cells or cell debris by filtration and centrifugation, followed by purification. In the purification, the culture products which have been separated from cells or cell debris can be subjected to conventional methods used to purify biologically-active proteins, for example, salting-out, dialysis, filtration, concentration, fractional precipitation, ion-exchange chromatography, gel filtration chromatography, adsorption chromatography, isoelectric focusing chromatography, hydrophobic chromatography, reversed phase chromatography, affinity chromatography, gel electrophoresis and isoelectric focusing gel electrophoresis which are used in combination, if necessary. The purified preparation of the present hedgehog protein can be concentrated and lyophilized into a liquid or solid form to meet to its final use. Immunoaffinity chromatographies using the monoclonal antibody described below do yield a high-purity preparation of the hedgehog protein with minimized costs and labors.

In the process of this invention for producing the hedgehog protein, the DNA expression step can also include a step of feeding or planting the transgenic animals or plants obtained by introducing the DNA which encodes the present hedgehog protein to animals other than humans or plants. After feeding or planting occasionally with desired stimuli, desired tissues, organs, bloods, milks, and/or body fluids of the resultants can be collected and subjected to the steps for purifying the hedgehog protein of this invention as mentioned above to obtain the present protein.

The monoclonal antibody of this invention includes the monoclonal antibodies in general which recognize the hedgehog protein of this invention, independently of its origins, sources, and classes. The monoclonal antibody of this invention can be obtained by using as an antigen the present hedgehog protein, other conventional hedgehog protein or antigenic fragment thereof, and more particularly, by preparing hybridoma cells of derived from an infinitely-proliferative of a mammal and an antibody-producing cell of a mammal that has been immunized with such an antigen, selecting clones of hybridoma capable of producing the monoclonal antibody of this invention, and culturing the clones *in vitro* or in *vivo*.

Proteins feasible as the antigens can be obtained through culturing of transformants introduced with a DNA encoding at least a partial amino acid sequence of SEQ ID NO:1, and the proteins are usually used after completely or partially purified. The antigenic fragments can be obtained by chemically or enzymatically digesting the completely or partially purified proteins or by chemical synthesis based on the amino acid sequence of SEQ ID NO:1, 2, or 3. Alternatively, the antigens can be obtained by using these techniques based on known hedgehog genes or proteins. Human Sonic hedgehog is useful as such known hedgehog.

Immunization of animals is conducted in conventional manner. For example, the antigens as described above can be injected alone or together with appropriate adjuvants into mammals through an intravenous, intradermal, subcutaneous or intraperitoneal route, and then the mammals can be fed for a prescribed time period. There is no limitation in types of the mammals, therefore any mammals can be used regardless of their types, sizes and genders, as far as one can obtain desired antigen-producing cells therefrom. Rodents such as rats, mice and hamsters are generally used, and among these the most desirable mammal can be chosen in respect to their compatibility with the infinitely-proliferative cells mentioned below. The dose of the antigen is generally set to about five to 500 μg/animal in total, which can be divided into two to five times inoculations with intervals of about one to two weeks, depending on the types and sizes of the mammals to be used. Three to five days after the final inoculation, the spleens are extracted and dispersed to obtain splenocytes as antibody-producing cells.

The antibody-producing cells obtained in this way can be then fused with infinitely-proliferative cells of mammalian origin to obtain cell-fusion products containing the objective hybridoma. Examples of the infinitely-proliferative cells usually used in this invention are cell lines of mouse myeloma origin such as P3/NSI/1-Ag4-1 cell, ATCC TIB-18; P3X63Ag8 cell, ATCC TIB-9; SP2/0-Ag14 cell, ATCC CRL-1581; and mutant strains thereof. The cell-fusion can be conducted in conventional manner using an electric pulse or a cell-fusion accelerator such as polyethylene glycol and Sendai virus. For example, the antibody-producing cells and the infinitely-proliferative cells of mammalian origin are co-suspended to give a ratio of about 1:1 to 1:10 in a cell fusion medium with such an accelerator and incubated at about 30 to 40°C for about one to five minutes. Although conventional media such as minimum essential medium (MEM), RPMI-1640 medium, and Iscove's modified Dulbecco's medium are feasible as cell fusion media, it is desirable

Example 4, and the method for detecting the hedgehog protein using the monoclonal antibody of this invention is explained by Examples 5 and 6. The following Examples can be diversified by the technical level in this field. In view of this, this invention should not be restricted to the Examples:

Example 1

Preparation of DNA

Example 1-1(a)

Preparation of total RNA

ARH-77 cells, ATCC CRL-1621, an established cell line derived from human plasma cell leukemia, were suspended in RPMI-1640 medium supplemented with 10%(v/v) fetal bovine serum and proliferated in usual manner at 37°C in a 5%(v/v) CO_2 incubator while scaling up the culture. After the cell density reached a desired level, the cells were collected. The cells were suspended in micro-centrifugal tubes with phosphate-beffered saline (hereinafter, abbreviated as "PBS") and centrifuged, and the supernatants were discarded; these treatments were repeated three times. Then the cells were placed in fresh micro-centrifugal tubes in an amount of 5×10^6 cells/tube, and "ULTRASPEC™ RNA", a total RNA isolation reagent commercialized by BIOTECX Laboratories, Inc., Houston, Texas, USA, was added to the tubes in a volume of 1.0 ml/tube before the cells were suspended. The suspensions were allowed to stand in ice-chilling conditions for 5 minutes, mixed with 1.2 ml/tube of a mixture of chloroform/"ULTRASPEC™ RNA" (1/5 by volume), shaken for 15 seconds, and allowed to stand in ice-chilling conditions for 5 minutes. Upper phase in the tubes formed by centrifugation was collected, mixed with the equal volume of 2-propanol, and allowed to stand in ice-chilling conditions for five minutes. The mixture was centrifuged, and the supernatant was discarded. The formed precipitate was washed twice with 75%(v/v) aqueous ethanol, dried up *in vacuo,* and dissolved in sterile distilled water, resulting in obtaining an aqueous solution containing total RNAs of ARH-77 cells. A small portion of the solution was examined for the absorbance at 260 nm to calculate an RNA content.

Example 1-1(b)

Preparation of first strand cDNA

Based on the nucleotide sequence of a mouse Desert hedgehog gene registered in "GenBank®", a nucleic acid database by National Institute of Health, USA, under the accession number "X76292", an oligonucleotide with the nucleotide sequence of 5'-GCCAGGGTGTGAGCAACAGT-3' (SEQ ID NO:12) was prepared in usual manner. In a micro-reaction tube, 2.5 pmol of the oligonucleotide and one microgram of total RNAs prepared by the method in Example 1-1(a) were placed, and sterile distilled water was added to the mixture to give a final volume of 15.5 µl. After the tube was allowed to stand at 70°C for ten minutes and under ice-chilling conditions for one minute, to the tube 2.5 µl of 10 x PCR buffer, 2.5 µl of 25 mM MgCl2, 1.0 µl of 10 mM dNTP mix, and 2.5 µl of 0.1 M DTT were added in this order. The tube was allowed to stand at 42°C for one minute. First strand cDNAs was synthesized by adding to the tube one microliter of "SUPERSCRIPT II RT", a reagent of reverse transcriptase commercialized by GIBCO BRL, Life Technologies, Inc., Rockville, Maryland, USA, and incubating the tube at 42°C for 50 minutes. After the mixture was heated to terminate the reaction at 70°C for 15 minutes and cooled to 37°C, the RNAs were degraded by incubating with admixed one microliter of RNase at 37°C for 30 minutes. Thereafter, from the reaction mixture, an aqueous solution containing purified first strand cDNAs in a volume of 50 µl was obtained by mixing with 120 µl of 6 M NaI and treating with "GlassMAX™", a DNA isolation matrix commercialized by GIBCO BRL, Life Technologies, Inc., Rockville, Maryland, USA, in accordance with the accompanying instructions.

Example 1-1(c)

Preparation of DNA fragment encoding the hedgehog protein and recombinant DNA

Ten-microliter portion of a solution of first strand cDNAs, obtained by the method in Example 1-1(b), was sampled in a micro-reaction tube and manipulated with "5' RACE SYSTEM, VERSION 2.0", a kit for a modified PCR method of 5' RACE, commercialized by GIBCO BRL Life Technologies, Inc., Rockville, Maryland, USA, in accordance with the accompanying instructions to add a poly(C)-tail to each of the 5'-termini of the cDNAs and amplify DNA fragments for the 5'-terminal regions. The sense primer used was "anchor primer" in the kit, and the antisense primer used was the oligonucleotide in Example 1-1(b). The thermal controls were as follows: an incubation at 94°C for one minute; 35

v) formamide, and 100 μg/ml denatured salmon sperm DNA, at 42°C for 1-2 hours, and subsequently immersed in a fresh pre-hybridization solution with an appropriate amount of the ^{32}P-labelled DNA fragment added as a probe and incubated at 42°C for 16-20 hours to effect hybridization. After the hybridization, the membranes were washed with 2 x SSC containing 0.1%(w/v) SDS at ambient temperature for 15 minutes and further washed with 0.2 x SSC containing 0.1%(w/v) SDS at a temperature moderately increasing from 37°C to 65 C until background radioactivity was adequately reduced. Thereafter the membranes were subjected to autoradiography. From a plaque which gave a positive signal, a phage clone was collected and amplified in usual manner, and from the amplified phage a DNA clone was collected. The DNA clone was sequenced by dideoxy method using primers prepared based on the vector's nucleotide sequence. The DNA clone contained a partial nucleotide sequence as shown with 5'-GTATCCATGGCTCTCCTG-3' (SEQ ID NO: 15). Compared with other known nucleotide sequences, the partial nucleotide sequence had a significant homology to a partial nucleotide sequence, containing translation initiation site, of a mouse Desert hedgehog gene registered in "GenBank®", a nucleic acid database by National Institute of Health, USA, under the accession number "X76292".

As sense and antisense primers for PCR, oligonucleotides with respective nucleotide sequences of 5'-GCCTC-GAGGTATCCATGGCTCTCCTG-3' (SEQ ID NO:16), which contains the above-determined partial nucleotide sequence, and 5'-GCGCGGCCGCTCAGCCGCCCGCCCGGAC-3' (SEQ ID NO:17), which is complementary to the sequence of nucleotides 532-548 of SEQ ID NO:7, were prepared in usual manner. As a template one microliter portion of cDNAs solution, obtained by the methods in Examples 1-1(a) and 1-1(b), was placed in a micro-reaction tube, then to which 3 μl of 10 x PCR buffer, 1.8 μl of 25 mM $MgCl_2$, 0.6 μl of 10 mM dNTP mix, appropriate amounts of the sense and antisense primers, and sterile distilled water were added to give a final volume of 30 μl. After 0.3 μl of 5 units/pl Taq DNA polymerase was added to thr tube, the mixture was subjected to an incubation at 94°C for three minutes, 35 cycles of incubations at 94°C for one minute, at 55°C for one minute, and at 72°C one minute, and finally an incubation at 72°C for 10 minutes, to effect PCR. The PCR products were subjected to 2%(w/v) agarose gel electrophoresis. From the gel, a gel portion containing an about 600 bp-DNA band, stained with ethidium bromide, was excised and treated with "SUPREC™-01", a DNA purification tube commercialized by Takara Shuzo Co., Ltd., Tokyo, Japan, to obtain 20 μl aqueous solution containing a DNA fragment.

A small portion of the DNA fragment solution was sampled and manipulated with "pT7BLUE CLONING KIT", a DNA cloning kit commercialized by Takara Shuzo Co., Ltd., Tokyo, Japan, in accordance with the accompanying instructions to ligate the DNA fragment with "pT7BLUE", the plasmid vecotor in the kit. After the ligation, a portion of the reaction mixture was introduced by usual transformation method into competent cells of *Escherichia coli* strain "JM101", commercialized by Takara Shuzo Co., Ltd., Tokyo, Japan, which were then inoculated on plates of L agar medium containing 50 μg/ml ampicillin and cultured at 37°C under standing conditions overnight. The formed colonies were respectively suspended in 10 μl aliquotes of sterile distilled water. Except for using the suspensions as respective templates, PCRs were conducted under the same conditions as in Example 1-1(c). Colonies which gave an about 600 bp-DNA band on agarose gel electrophoresis were respectively inoculated to aliquotes of L broth medium containing 50 μg/ml ampicillin and cultured at 37°C under shaking conditions overnight. From the resulting cultures recombinant DNAs were collected by conventional alkali-SDS method. The recombinant DNAs were sequenced by dideoxy method. The DNA fragment in the recombinant DNAs contained the nucleotide sequence of SEQ ID NO:8, which can encode the amino acid sequence shown along with the nucleotide sequence.

The nucletotide sequence of SEQ ID NO:8 was compared with the nucleotide sequence of SEQ ID NO:7, determined in Example 1-1. The sequence of nucleotides 1-548 of SEQ ID NO:7 completely matched with the sequence of nucleotides 55-602 of SEQ ID NO:8. The results of this comparison and the comparison with the above-mentioned nucleotide sequence of a mouse Desert hedgehog gene revealed that: the nucleotide sequence of SEQ ID NO:8 encodes N-terminal region of a precursor form of a human Desert hedgehog protein; the sequence of nucleotides 7-72 of SEQ ID NO:8 encodes a signal peptide in a precursor form of the hedgehog protein; and the sequence of nucleotides 73-600 of SEQ ID NO:8 encodes a mature form of the hedgehog protein which contains the amino acid sequence of SEQ ID NO:1.

Example 1-3

Preparation of DNA fragment and recombinant DNA encoding the hedgehog protein

As sense and antisense primers for PCR, oligonucleotides with respective nucleotide sequences of 5'-CGTGTCG-GTCAAAGCTGATA-3' (SEQ ID NO:18) and 5'-ATGCATTCCAGTCGGCTGGA-3' (SEQ ID NO:19) were prepared in usual manner; the former sequence was identical to the sequence of nucleotides 501-520 of SEQ ID NO:7, and the latter sequence was based on the nucleotide sequence registered in "GenBank®", a nucleic acid database by National Institute of Health, USA, under the accession number "AA064660", which is of a human cDNA fragment similar to a 3'-terminal sequence for a mouse Desert hedgehog protein in a precursor form. As a template one microliter portion of cDNAs solution, obtained by the methods in Examples 1-1(a) and 1-1(b), was placed in a micro-reaction tube, then

obtain 20 µl aqueous solution containing a DNA fragment.

A portion of the DNA fragment solution was sampled and manipulated with "pT7BLUE CLONING KIT", a DNA cloning kit commercialized by Takara Shuzo Co., Ltd., Tokyo, Japan, in accordance with the accompanying instructions to ligate the DNA fragment with "pT7BLUE", the plasmid vector in the kit. After the ligation, a portion of the reaction mixture was introduced by usual transformation method into competent cells of *Escherichia coli* strain "JM101", commercialized by Takara Shuzo Co., Ltd., Tokyo, Japan, which were then inoculated on plates of L agar medium containing 50 µg/ml ampicillin and cultured at 37°C under standing conditions overnight. The formed colonies were respectively suspended in 10 µl aliquotes of sterile distilled water. Except for using the suspensions as respective templates and using the sense and antisense primers in the third step PCR, PCRs were conducted under the same conditions as in Example 1-1(c). Colonies which gave an about 750 bp-DNA band on agarose gel electrophoresis were respectively inoculated to aliquotes of L broth medium containing 50 µg/ml ampicillin and cultured at 37°C under shaking conditions overnight. From the resulting cultures recombinant DNAs were collected by alkali-SDS method. The recombinant DNAs were sequenced by dideoxy method. The DNA fragment in the recombinant DNAs contained the nucleotide sequence of SEQ ID NO:10, which can encode the amino acid sequence shown along with the nucleotide sequence.

The nucleotide sequence of SEQ ID NO:10 was compared with the nucleotide sequence of SEQ ID NO:9, determined in Example 1-3. The sequence of nucleotides 1-152 of SEQ ID NO:10 completely matched with the sequence of nucleotides 424-575 of SEQ ID NO:9. The results of this comparison and the comparison with the above-mentioned nucleotide sequence of a mouse Desert hedgehog gene revealed that the nucleotide sequence of SEQ ID NO:10 encodes a region containing the C-terminus of a precursor form of a human Desert hedgehog protein.

As described in Examples 1-1 to 1-4, the nucleotide sequences of SEQ ID NOs:7-10, determined in these Examples, were proved to be overlapping nucleotide sequences one another which partially encode a precursor form of a human Desert hedgehog protein; and the precursor protein can be wholly encoded by a DNA containing the nucleotide sequence of SEQ ID NO:6. In addition, these results elucidated that: a human Desert hedgehog protein can be in a precursor form which contains the amino acid sequence of SEQ ID NO:2 or 3 or in a mature form which contains the amino acid sequence of SEQ ID NO:1; such precursor protein can be encoded by a DNA containing the nucleotide sequence of SEQ ID NO:5 or 6, respectively; and such mature protein can be encoded by a DNA containing the nucleotide sequence of SEQ ID NO:4.

Example 2

Preparation of transformant

Based on the nucleotide sequence determined in Example 1-1(c), which encodes a precursor form of a human Desert hedgehog protein, oligonucleotides with respective nucleotide sequences of 5'-CCCGGGAATTCATTGCGGGCCGGGCCGGGGGCCG-3' (SEQ ID NO:27) as a sense primer and 5'-ACGATGAATTCTCAGCCGCCCGCCCGGACCGCCA-3' (SEQ ID NO:28) as an antisense primer were prepared in usual manner. PCR was conducted under the same conditions as in Example 1-1(c) except for using the recombinant DNA "pHuDHH/#20" as a template, obtained by the method in Example 1-1(c), and the above sense and antisense primers. An about 600 bp-DNA amplified in this PCR was purified by 2%(w/v) agarose gel electrophoresis and treating with "SUPREC™-01", a DNA purification tube commercialized by Takara Shuzo Co., Ltd., Tokyo, Japan, in accordance with the accompanying instruction, to obtain 20 µl aqueous DNA solution. Two microliters portion of the DNA solution was sampled and subjected to a ligation reaction using T4 DNA ligase with "pCR™II", a plasmid vector for TA cloning commercialized by Invitrogen Corporation, San Diego, USA. A portion of the reaction mixture was introduced by usual transformation method into competent cells of *Escherichia coli* strain "TDP10F'", commercialized by Invitrogen Corporation, San Diego, USA, which were then inoculated on plates of L agar medium containing 50 µg/ml ampicillin and 50 µg/ml 5-bromo-4-chrolo-3-indolyl-β-D-galactoside and cultured at 37°C under standing conditions. A white colony formed was inoculated to an aliquote of L broth medium containing 50 µg/ml ampicillin and cultured at 37°C under shaking conditions overnight. From the resulting culture, a recombinant DNA was collected by alkali-SDS method. The recombinant DNA was treated with restriction enzyme EcoRI and subjected to 2%(w/v) agarose gel electrophoresis, on which an about 600 bp-DNA was separated, and it was then purified with "SUPREC™-01", a DNA purification tube commercialized by Takara Shuzo Co., Ltd, Tokyo, Japan.

A portion of the purified DNA solution was sampled and subjected to a ligation reaction in usual manner using T4 DNA ligase with plasmid vector "pGEX-2T", commercialized by Pharmacia Biotech, Inc., Uppsala, Sweden, which had been cleaved with EcoRI and dephosphorylated prior to use. A portion of the ligation reaction mixture was introduced by usual transformation method into competent cells prepared by applying the method in *DNA cloning,* Vol.1, edited by D. M. Glover, published by IRL press limited, Oxford, England (1985), pp.109-136, to *Escherichia coli* "BL21" strain, commercialized by Pharmacia Biotech, Inc., Uppsala, Sweden, which were then inoculated to plates of L agar medium containing 50 µg/ml ampicillin and cultured at 37°C under standing conditions overnight. A colony formed was inocu-

Example 4

Production of monoclonal antibody

Example 4-1

Preparation of immunogen

Example 4-1(a)

Preparation of transformant introduced with DNA that encodes immunogen

A549 cells, ATCC CCL-185, an established cell line derived from a human lung carcinoma, were suspended in RPMI-1640 medium (pH 7.2) supplemented with 10%(v/v) fetal bovine serum and proliferated in usual manner at 37°C in a 5%(v/v) CO_2 incubator while scaling up the culture. After the cell density reached a desired level, proliferated cells were collected. The cells were manipulated with "ULTRASPEC™ RNA", similarly as in Example 1-1(a), to obtain an aqueous solution containing total RNAs of A549 cells. By applying usual RT-PCR method to the total RNAs, a DNA fragment encoding a mature form of a human Sonic hedgehog protein was amplified. As the sense and antisense primers in this RT-PCR, oligonucleotides with respective nucleotide sequences of 5'-CCCGGGAATTCATTGCGGAC-CGGGCAGGGGGTT-3' (SEQ ID NO:30) and 5'-ACGATGAATTCTCAGCCTCCCGATTTGGCCGC-3' (SEQ ID NO:31), prepared in usual manner based on the nucleotide sequence of a human Sonic hedgehog gene, reported by V. Marigo et al. and registered in "GenBank®", a nucleic acid database by National Institute of Health, USA, under the accession number "L38518", were used. The amplified DNA fragment was collected by treating the RT-PCR products with "SUPREC™-01", as in Example 1-1(c). Similarly as in Example 1-1(c), the DNA fragment was ligated with plasmid vector "pCR™II" and introduced into *Escherichia coli* "TDP10F'" strain, the obtained transformant was cultured, and from the resulting culture a recombinant DNA was collected by alkali-SDS method. The recombinant DNA was sequenced by dideoxy method, confirming that it contained the nucleotide sequence of SEQ ID NO:11, encoding a human Sonic hedgehog protein in a mature form.

Similarly as in Example 2, an aliquot of the recombinant DNA was cleaved with restriction enzyme *Eco*RI to form an about 600 bp-DNA, which was then collected by treating with "SUPREC™-01", ligated with plasmid vector "pGEX-2T", and introduced into *Escherichia coli* "BL21" strain. The obtained transformant was cultured, and from the resulting culture a recombinant DNA was collected by alkali-SDS method. The recombinant DNA was sequenced by dideoxy method, confirming that it contained a DNA with the nucleotide sequence of SEQ ID NO:11 and a termination codon, which are respectively located in the downstream and further downstream of a structural gene of glutathione S-transferase in the same frame under the regulation of Tac promotor. The recombinant DNA and the transformant with the recombinant DNA introduced, thus obtained, were named "pHuSHH/pGEX-2T/#3-1" and "TAL#3-1/HuSHH", respectively.

Example 4-1(b)

Preparation of immunogen using transformant

Similarly as in Example 3, the transformant "TAL#3-1/HuSHH" obtained by the method in Example 4-1(a) was cultured, the proliferated cells were collected from the culture, and a supernatant of the cell-disruptant was obtained. By applying the methods using "GLUTATHIONE SEPHAROSE 4B BEADS", thrombin, "ANTITHROMBIN AGAROSE" and "HEPARIN AGAROSE" in Example 3 to the supernatant, an aqueous solution containing a protein derived from "TAL#3-1/HuSHH" was obtained, and analyzed by SDS-PAGE; a main band was observed at a position corresponding to a molecular weight of 22,000±2,000. The molecular weight of a mature form of a human Sonic hedgehog protein which has the amino acid sequence shown along with SEQ ID NO:11 is calculated to be 19,747. According to this Example, the objective protein is usually generated in a form with a peptide as shown by Gly-Ser-Pro-Gly-Ile-His- (SEQ ID NO:29) added to the N-terminus and collected. These indicate that the protein obtained in this Example is a human Sonic hedgehog protein with a satisfactory purity. Thus, a purified preparation of a human Sonic hedgehog protein as an immunogen was obtained.

v) "BLOCK ACE™", and 0.05%(v/v) "TWEEN 20"; washed with PBS containing 0.1%(v/v) "TWEEN 20"; and color-developed by using "ECL™ KIT", a kit for color development commercialized by Amersham International plc, Buckinghamshire, UK. The molecular weight markers used were "SDS-PAGE STANDARDS, LOW RANGE", containing six proteins having distinctive molecular weights of 14,400-97,400 daltons, commercialized by Bio-rad Laboratories Inc., Richmond, USA. The results are in FIG.2.

In FIG.2, on Lane 1, the band corresponding to a molecular weight of 22,000±2,000 is of the hedgehog protein of this invention, and the other band, corresponding to a molecular weight of 18,000±2,000, is of the degradation product of the hedgehog protein formed during the process in Example 3. In FIG.2, on Lane 2, the band corresponding to a molecular weight of 22,000±2,000 is of a human Sonic hedgehog protein, obtained by the method in Example 4-1(b).

Another Western blotting which was conducted in the same manner as above except for using a monoclonal antibody "SH2-21mAb", obtained by the method in Example 4-3, in place of the monoclonal antibody "SH2-3mAb", giving similar results as above. These results indicate that the monoclonal antibodies, according to this invention, well recognized not only a human Sonic hedgehog protein but also a human Desert hedgehog protein.

Example 6

Enzyme-immunoassay

Monoclonal antibodies "SH2-3mAb" and "SH2-260mAb", obtained by the method in Example 4-3, were co-diluted in PBS to give a concentration of 10 μg/ml each, the resulting solution was distributed to wells of 96-well microplates in a volume of 100 μl/well. The microplates were incubated at ambient temperature. From the microplates the solution was removed, and PBS containing 1%(w/v) bovine serum albumin was distributed to the wells in a volume of 200 μl/well. Then the microplates were allowed to stand at 4°C overnight. In parallel, a human Desert hedgehog protein, obtained by the method in Example 3, and a human Sonic hedgehog protein, obtained by the method in Example 4-1(b), were separately diluted with PBS to give desired different concentrations. After removing the solution from the microplates, and the respective hedgehog protein solutions were added to the wells and reacted at ambient temperature for one hour. The microplates were washed with PBS containing 0.05%(v/v) "TWEEN 20", and added with a rabbit anti-hedgehog protein antiserum 500-fold diluted with PBS in a volume of 100 μl/well. The antiserum used in this Example was obtained by immunizing rabbits with a human Sonic hedgehog protein, obtained by the method in Example 4-1(b), and collecting serum from the rabbits in usual manner.

After the reaction with the antiserum, the microplates were washed with PBS containing 0.05%(v/v) "TWEEN 20" and added with a horseradish peroxidase-labelled donkey anti-rabbit immunoglobulin antibody, commercialized by Amersham International plc, Buckinghamshire, UK, which had been 1000-fold diluted with PBS, followed by allowing the microplates to stand at ambient temperature for one hour. The microplates were washed with PBS containing 0.05%(v/v) "TWEEN 20". Thereafter, in usual manner, a mixture solution of o-phenylene diamine as a substrate and H_2O_2 was added to the wells in a volume of 100 μl/well followed by an incubation at ambient temperature for 15 minutes to effect enzyme reaction, and the reaction was terminated by 2N H_2SO_4 added. Intensities of colors in the wells developed by the reaction were estimated by measuring the absorbance at 492 nm. The results are in FIG.3.

The results in FIG.3. indicate that the method for detecting, according to this invention, well detected not only a human Sonic hedgehog protein but also a human Desert hedgehog protein.

As described above, this invention was established based on the finding of a novel hedgehog protein, i.e., a Desert hedgehog protein of human origin. The hedgehog protein of this invention is useful in establishment of a hybridoma capable of producing a monoclonal antibody that recognizes the protein. The hedgehog protein of this invention has efficacy in treatment and prevention of susceptive diseases to the hedgehog protein. The monoclonal antibody is useful in purification and detection of human Desert hedgehog protein because the antibody recognizes the hedgehog protein. The monoclonal antibody has efficacy in treatment, prevention, and diagnosis of diseases relating to excessive production of the hedgehog protein in living bodies. In addition to these effectiveness, the protein, DNA, and monoclonal antibody of this invention are extremely useful in elucidation of the process of exhibiting hereditary morphological abnormalities in humans. The process of this invention does satisfactorily produce the hedgehog protein.

This invention, which exhibits these remarkable effects, would be very significant and contributive to the art.

While there has been described what is at present considered to be the preferred embodiments of the present invention, it will be understood the various modifications may be made therein, and it is intended to cover in the appended claims all such modifications as fall within the true spirits and scope of the invention.

```
          35                  40                  45

Ser Glu Arg Phe Arg Asp Leu Val Pro Asn Tyr Asn Pro Asp Ile Ile
     50                  55                  60

Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp Arg Leu Met Thr Glu Arg
 65                  70                  75                  80

Cys Lys Glu Arg Val Asn Ala Leu Ala Ile Ala Val Met Asn Met Trp
                 85                  90                  95

Pro Gly Val Arg Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His
            100                 105                 110

His Ala Gln Asp Ser Leu His Tyr Glu Gly Arg Ala Leu Asp Ile Thr
        115                 120                 125

Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala
    130                 135                 140

Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Arg Asn His Ile
145                 150                 155                 160

His Val Ser Val Lys Ala Asp Asn Ser Leu Ala Val Arg Ala Gly Gly
                165                 170                 175
```

(3) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 374 amino acids
(B) TYPE: amino acid
(C) strandedness: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: mat peptide
(B) LOCATION: 1..176
(C) IDENTIFICATION METHOD: S

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Cys Gly Pro Gly Arg Gly Pro Val Gly Arg Arg Arg Tyr Ala Arg Lys
 1               5                  10                  15

Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe Val Pro Gly Val Pro Glu
            20                  25                  30

Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu Gly Arg Val Ala Arg Gly
        35                  40                  45

Ser Glu Arg Phe Arg Asp Leu Val Pro Asn Tyr Asn Pro Asp Ile Ile
     50                  55                  60

Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp Arg Leu Met Thr Glu Arg
 65                  70                  75                  80
```

(4) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 396 amino acids
(B) TYPE: amino acid
(C) strandedness: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: sig peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: S
(A) NAME/KEY: mat peptide
(B) LOCATION: 1..176
(C) IDENTIFICATION METHOD: S

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
Met Ala Leu Leu Thr Asn Leu Leu Pro Leu Cys Cys Leu Ala Leu Leu
        -20                 -15                 -10

Ala Leu Pro Ala Gln Ser Cys Gly Pro Gly Arg Gly Pro Val Gly Arg
     -5                  1                   5                  10

Arg Arg Tyr Ala Arg Lys Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe
                15                  20                  25

Val Pro Gly Val Pro Glu Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu
            30                  35                  40

Gly Arg Val Ala Arg Gly Ser Glu Arg Phe Arg Asp Leu Val Pro Asn
        45                  50                  55

Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp
     60                  65                  70

Arg Leu Met Thr Glu Arg Cys Lys Glu Arg Val Asn Ala Leu Ala Ile
 75                  80                  85                  90

Ala Val Met Asn Met Trp Pro Gly Val Arg Leu Arg Val Thr Glu Gly
                95                  100                 105

Trp Asp Glu Asp Gly His His Ala Gln Asp Ser Leu His Tyr Glu Gly
            110                 115                 120

Arg Ala Leu Asp Ile Thr Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly
        125                 130                 135

Leu Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr
    140                 145                 150

Glu Ser Arg Asn His Ile His Val Ser Val Lys Ala Asp Asn Ser Leu
155                 160                 165                 170

Ala Val Arg Ala Gly Gly Cys Phe Pro Gly Asn Ala Thr Val Arg Leu
```

```
          1               5                  10                  15

CAG CTC GTG CCG CTA CTC TAC AAG CAA TTT GTG CCC GGC GTG CCA GAG   96
Gln Leu Val Pro Leu Leu Tyr Lys Gln Phe Val Pro Gly Val Pro Glu
            20                  25                  30

CGG ACC CTG GGC GCC AGT GGG CCA GCG GAG GGG AGG GTG GCA AGG GGC  144
Arg Thr Leu Gly Ala Ser Gly Pro Ala Glu Gly Arg Val Ala Arg Gly
        35                  40                  45

TCC GAG CGC TTC CGG GAC CTC GTG CCC AAC TAC AAC CCC GAC ATC ATC  192
Ser Glu Arg Phe Arg Asp Leu Val Pro Asn Tyr Asn Pro Asp Ile Ile
    50                  55                  60

TTC AAG GAT GAG GAG AAC AGT GGA GCC GAC CGC CTG ATG ACC GAA CGT  240
Phe Lys Asp Glu Glu Asn Ser Gly Ala Asp Arg Leu Met Thr Glu Arg
65                  70                  75                  80

TGT AAG GAA CGG GTG AAC GCT TTG GCC ATT GCC GTG ATG AAC ATG TGG  288
Cys Lys Glu Arg Val Asn Ala Leu Ala Ile Ala Val Met Asn Met Trp
                85                  90                  95

CCC GGA GTG CGC CTA CGA GTG ACT GAG GGC TGG GAC GAG GAC GGC CAC  336
Pro Gly Val Arg Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His
            100                 105                 110

CAC GCT CAG GAT TCA CTC CAC TAC GAA GGC CGT GCT TTG GAC ATC ACT  384
His Ala Gln Asp Ser Leu His Tyr Glu Gly Arg Ala Leu Asp Ile Thr
        115                 120                 125

ACG TCT GAC CGC GAC CGC AAC AAG TAT GGG TTG CTG GCG CGC CTC GCA  432
Thr Ser Asp Arg Asp Arg Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala
    130                 135                 140

GTG GAA GCC GGC TTC GAC TGG GTC TAC TAC GAG TCC CGC AAC CAC ATC  480
Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Arg Asn His Ile
145                 150                 155                 160

CAC GTG TCG GTC AAA GCT GAT AAC TCA CTG GCG GTC CGG GCG GGC GGC  528
His Val Ser Val Lys Ala Asp Asn Ser Leu Ala Val Arg Ala Gly Gly
                165                 170                 175
```

(6) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1122 base pairs
(B) TYPE: nucleic acid
(C) strandedness: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: mat peptide
(B) LOCATION: 1..528
(C) IDENTIFICATION METHOD: S

```
TTG CAG CGC CGG GCT TCA TTT GTG GCT GTG GAG ACC GAG TGG CCT CCA  720
Leu Gln Arg Arg Ala Ser Phe Val Ala Val Glu Thr Glu Trp Pro Pro
225                 230                 235                 240

CGC AAA CTG TTG CTC ACG CCC TGG CAC CTG GTG TTT GCC GCT CGA GGG  768
Arg Lys Leu Leu Leu Thr Pro Trp His Leu Val Phe Ala Ala Arg Gly
                245                 250                 255

CCG GCG CCC GCG CCA GGC GAC TTT GCA CCG GTG TTC GCG CGC CGG CTA  816
Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro Val Phe Ala Arg Arg Leu
            260                 265                 270

CGC GCT GGG GAC TCG GTG CTG GCG CCC GGC GGG GAT GCG CTT CGG CCA  864
Arg Ala Gly Asp Ser Val Leu Ala Pro Gly Gly Asp Ala Leu Arg Pro
        275                 280                 285

GCG CGC GTG GCC CGT GTG GCG CGG GAG GAA GCC GTG GGC GTG TTC GCG  912
Ala Arg Val Ala Arg Val Ala Arg Glu Glu Ala Val Gly Val Phe Ala
    290                 295                 300

CCG CTC ACC GCG CAC GGG ACG CTG CTG GTG AAC GAT GTC CTG GCC TCT  960
Pro Leu Thr Ala His Gly Thr Leu Leu Val Asn Asp Val Leu Ala Ser
305                 310                 315                 320

TGC TAC GCG GTT CTG GAG AGT CAC CAG TGG GCG CAC CGC GCT TTT GCC 1008
Cys Tyr Ala Val Leu Glu Ser His Gln Trp Ala His Arg Ala Phe Ala
                325                 330                 335

CCC TTG AGA CTG CTG CAC GCG CTA GGG GCG CTG CTC CCC GGC GGG GCC 1056
Pro Leu Arg Leu Leu His Ala Leu Gly Ala Leu Leu Pro Gly Gly Ala
            340                 345                 350

GTC CAG CCG ACT GGC ATG CAT TGG TAC TCT CGG CTC CTC TAC CGC TTA 1104
Val Gln Pro Thr Gly Met His Trp Tyr Ser Arg Leu Leu Tyr Arg Leu
        355                 360                 365

GCG GAG GAG CTA CTG GGC                                         1122
Ala Glu Glu Leu Leu Gly
    370
```

(7) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1188 base pairs
(B) TYPE: nucleic acid
(C) strandedness: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: sig peptide
(B) LOCATION: 1..66
(C) IDENTIFICATION METHOD: S
(A) NAME/KEY: mat peptide
(B) LOCATION: 67..594
(C) IDENTIFICATION METHOD: S

```
GTT TTG ACG GCC GAT GCG TCA GGC CGG GTG GTG CCC ACG CCG GTG CTG  720
Val Leu Thr Ala Asp Ala Ser Gly Arg Val Val Pro Thr Pro Val Leu
        205                 210                 215

CTC TTC CTG GAC CGG GAC TTG CAG CGC CGG GCT TCA TTT GTG GCT GTG  768
Leu Phe Leu Asp Arg Asp Leu Gln Arg Arg Ala Ser Phe Val Ala Val
    220                 225                 230

GAG ACC GAG TGG CCT CCA CGC AAA CTG TTG CTC ACG CCC TGG CAC CTG  816
Glu Thr Glu Trp Pro Pro Arg Lys Leu Leu Leu Thr Pro Trp His Leu
235                 240                 245                 250

GTG TTT GCC GCT CGA GGG CCG GCG CCC GCG CCA GGC GAC TTT GCA CCG  864
Val Phe Ala Ala Arg Gly Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro
                255                 260                 265

GTG TTC GCG CGC CGG CTA CGC GCT GGG GAC TCG GTG CTG GCG CCC GGC  912
Val Phe Ala Arg Arg Leu Arg Ala Gly Asp Ser Val Leu Ala Pro Gly
            270                 275                 280

GGG GAT GCG CTT CGG CCA GCG CGC GTG GCC CGT GTG GCG CGG GAG GAA  960
Gly Asp Ala Leu Arg Pro Ala Arg Val Ala Arg Val Ala Arg Glu Glu
        285                 290                 295

GCC GTG GGC GTG TTC GCG CCG CTC ACC GCG CAC GGG ACG CTG CTG GTG 1008
Ala Val Gly Val Phe Ala Pro Leu Thr Ala His Gly Thr Leu Leu Val
    300                 305                 310

AAC GAT GTC CTG GCC TCT TGC TAC GCG GTT CTG GAG AGT CAC CAG TGG 1056
Asn Asp Val Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp
315                 320                 325                 330

GCG CAC CGC GCT TTT GCC CCC TTG AGA CTG CTG CAC GCG CTA GGG GCG 1104
Ala His Arg Ala Phe Ala Pro Leu Arg Leu Leu His Ala Leu Gly Ala
                335                 340                 345

CTG CTC CCC GGC GGG GCC GTC CAG CCG ACT GGC ATG CAT TGG TAC TCT 1152
Leu Leu Pro Gly Gly Ala Val Gln Pro Thr Gly Met His Trp Tyr Ser
            350                 355                 360

CGG CTC CTC TAC CGC TTA GCG GAG GAG CTA CTG GGC                 1188
Arg Leu Leu Tyr Arg Leu Ala Glu Glu Leu Leu Gly
        365                 370
```

(8) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 548 base pairs
(B) TYPE: nucleic acid
(C) strandedness: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(B) INDIVIDUAL ISOLATE: ARH-77, ATCC CRL-1621

(9) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 602 base pairs
(B) TYPE: nucleic acid
(C) strandedness: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(B) INDIVIDUAL ISOLATE: ARH-77, ATCC CRL-1621

(ix) FEATURE:
(A) NAME/KEY: 5'UTR
(B) LOCATION: 1..6
(C) IDENTIFICATION METHOD: S
(A) NAME/KEY: sig peptide
(B) LOCATION: 7..72
(C) IDENTIFICATION METHOD: S
(A) NAME/KEY: mat peptide
(B) LOCATION: 73..600
(C) IDENTIFICATION METHOD: S

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
GTATCC ATG GCT CTC CTG ACC AAT CTA CTG CCC CTG TGC TGC TTG GCA     48
       Met Ala Leu Leu Thr Asn Leu Leu Pro Leu Cys Cys Leu Ala
                -20                 -15                 -10

CTT CTG GCG CTG CCA GCC CAG AGC TGC GGG CCG GGC CGG GGG CCG GTT   96
Leu Leu Ala Leu Pro Ala Gln Ser Cys Gly Pro Gly Arg Gly Pro Val
            -5                   1                   5

GGC CGG CGC CGC TAT GCG CGC AAG CAG CTC GTG CCG CTA CTC TAC AAG  144
Gly Arg Arg Arg Tyr Ala Arg Lys Gln Leu Val Pro Leu Leu Tyr Lys
     10                  15                  20

CAA TTT GTG CCC GGC GTG CCA GAG CGG ACC CTG GGC GCC AGT GGG CCA  192
Gln Phe Val Pro Gly Val Pro Glu Arg Thr Leu Gly Ala Ser Gly Pro
 25                  30                  35                  40

GCG GAG GGG AGG GTG GCA AGG GGC TCC GAG CGC TTC CGG GAC CTC GTG  240
Ala Glu Gly Arg Val Ala Arg Gly Ser Glu Arg Phe Arg Asp Leu Val
                45                  50                  55

CCC AAC TAC AAC CCC GAC ATC ATC TTC AAG GAT GAG GAG AAC AGT GGA  288
Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Ser Gly
            60                  65                  70

GCC GAC CGC CTG ATG ACC GAA CGT TGT AAG GAA CGG GTG AAC GCT TTG  336
Ala Asp Arg Leu Met Thr Glu Arg Cys Lys Glu Arg Val Asn Ala Leu
        75                  80                  85

GCC ATT GCC GTG ATG AAC ATG TGG CCC GGA GTG CGC CTA CGA GTG ACT  384
Ala Ile Ala Val Met Asn Met Trp Pro Gly Val Arg Leu Arg Val Thr
```

```
CGC AAA CTG TTG CTC ACG CCC TGG CAC CTG GTG TTT GCC GCT CCA GGG  286
Arg Lys Leu Leu Leu Thr Pro Trp His Leu Val Phe Ala Ala Arg Gly
 80                  85                  90                  95

CCG GCG CCC GCG CCA GGC GAC TTT GCA CCG GTG TTC GCG CGC CGG CTA  334
Pro Ala Pro Ala Pro Gly Asp Phe Ala Pro Val Phe Ala Arg Arg Leu
                100                 105                 110

CGC GCT GGG GAC TCG GTG CTG GCG CCC GGC GGG GAT GCG CTT CGG CCA  382
Arg Ala Gly Asp Ser Val Leu Ala Pro Gly Gly Asp Ala Leu Arg Pro
            115                 120                 125

GCG CGC GTG GCC CGT GTG GCG CGG GAG GAA GCC GTG GGC GTG TTC GCG  430
Ala Arg Val Ala Arg Val Ala Arg Glu Glu Ala Val Gly Val Phe Ala
        130                 135                 140

CCG CTC ACC GCG CAC GGG ACG CTG CTG GTG AAC GAT GTC CTG GCC TCT  478
Pro Leu Thr Ala His Gly Thr Leu Leu Val Asn Asp Val Leu Ala Ser
    145                 150                 155

TGC TAC GCG GTT CTG GAG AGT CAC CAG TGG GCG CAC CGC GCT TTT GCC  526
Cys Tyr Ala Val Leu Glu Ser His Gln Trp Ala His Arg Ala Phe Ala
160                 165                 170                 175

CCC TTG AGA CTG CTG CAC GCG CTA GGG GCG CTG CTC CCC GGC GGG GCC  574
Pro Leu Arg Leu Leu His Ala Leu Gly Ala Leu Leu Pro Gly Gly Ala
                180                 185                 190

G                                                               575
```

(11) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 230 base pairs
(B) TYPE: nucleic acid
(C) strandedness: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(B) INDIVIDUAL ISOLATE: ARH-77, ATCC CRL-1621

(ix) FEATURE:
(A) NAME/KEY: 3'UTR
(B) LOCATION: 218..230
(C) IDENTIFICATION METHOD: S

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
G TTC GCG CCG CTC ACC GCG CAC GGG ACG CTG CTG GTG AAC GAT GTC     46
  Phe Ala Pro Leu Thr Ala His Gly Thr Leu Leu Val Asn Asp Val
    1               5                  10                  15

CTG GCC TCT TGC TAC GCG GTT CTG GAG AGT CAC CAG TGG GCG CAC CGC   94
Leu Ala Ser Cys Tyr Ala Val Leu Glu Ser His Gln Trp Ala His Arg
```

```
                    85                  90                  95

GTG AAA CTG CGG GTG ACC GAG GGC TGG GAC GAA GAT GGC CAC CAC TCA  336
Val Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His His Ser
            100                 105                 110

GAG GAG TCT CTG CAC TAC GAG GGC CGC GCA GTG GAC ATC ACC ACG TCT  384
Glu Glu Ser Leu His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser
        115                 120                 125

GAC CGC GAC CGC AGC AAG TAC GGC ATG CTG GCC CGC CTG GCG GTG GAG  432
Asp Arg Asp Arg Ser Lys Tyr Gly Met Leu Ala Arg Leu Ala Val Glu
    130                 135                 140

GCC GGC TTC GAC TGG GTG TAC TAC GAG TCC AAG GCA CAT ATC CAC TGC  480
Ala Gly Phe Asp Trp Val Tyr Tyr Glu Ser Lys Ala His Ile His Cys
145                 150                 155                 160

TCG GTG AAA GCA GAG AAC TCG GTG GCG GCC AAA TCG GGA GGC          522
Ser Val Lys Ala Glu Asn Ser Val Ala Ala Lys Ser Gly Gly
                165                 170                 174
```

(13) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GCCAGGGTGT GAGCAACAGT 20

(14) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TGTGCTGCTT GGCACTCTTG 20

(15) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CCGTGGCATT TCCCGGAAAG 20

ATGCATTCCA GTCGGCTGGA 20

(21) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AAGGATCCGT CGACAAGCTT AATACGACGA ATTCTGGAGT TTTTTTTTTT TTTTTT 56

(22) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GGCTTCGACT GGGTCTACTA 20

(23) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

AAGGATCCGT CGACAAG 17

(24) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ATGCGCTTCG GCCAGCG 17

(25) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Gly Ser Pro Gly Ile His
1 5

(31) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CCCGGGAATT CATTGCGGAC CGGGCAGGGG GTT 33

(32) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) strandedness: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

ACGATGAATT CTCAGCCTCC CGATTTGGCC GC 32

Claims

1. A Desert hedgehog protein of human origin.

2. The hedgehog protein of claim 1, which contains a part or the whole of the amino acid sequence of SEQ ID NO:1.

3. The hedgehog protein of claim 1 or claim 2, which contains a part or the whole of the amino acid sequence of SEQ ID NO:2.

4. The hedgehog protein of claim 1 or claim 2, which contains a part or the whole of the amino acid sequence of SEQ ID NO:3.

5. The hedgehog protein of any one of claims 1 to 4, which originates from a human cell.

6. The hedgehog protein of any one of claims 1 to 5, which originates from established human cell line ARH-77, ATCC CRL-1621.

7. A DNA which encodes the hedgehog protein of claim 1.

8. The DNA of claim 7, which contains a part or the whole of either the nucleotide sequence of SEQ ID NO:4 or its complementary nucleotide sequence.

9. The DNA of claim 7 or claim 8, which contains a part or the whole of either the nucleotide sequence of SEQ ID NO:5 or its complementary nucleotide sequence.

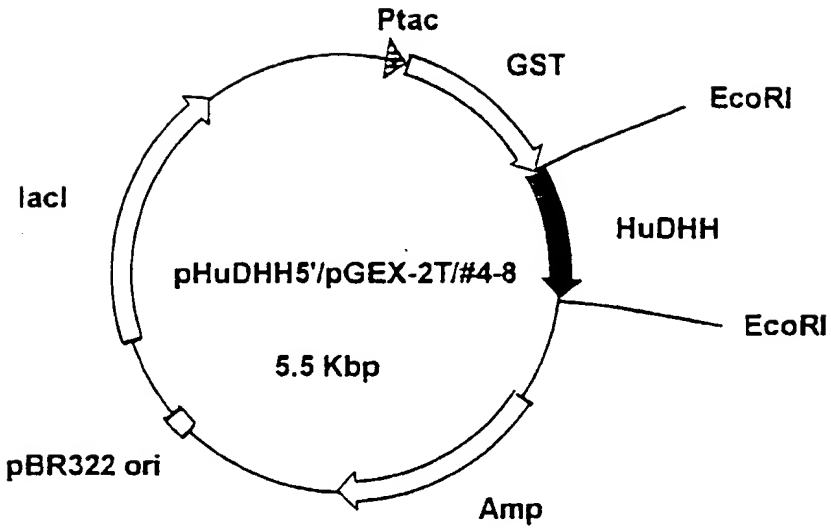
Ptac
GST
EcoRI
HuDHH
EcoRI
Amp
pBR322 ori
lacI
pHuDHH5'/pGEX-2T/#4-8
5.5 Kbp

FIG.1

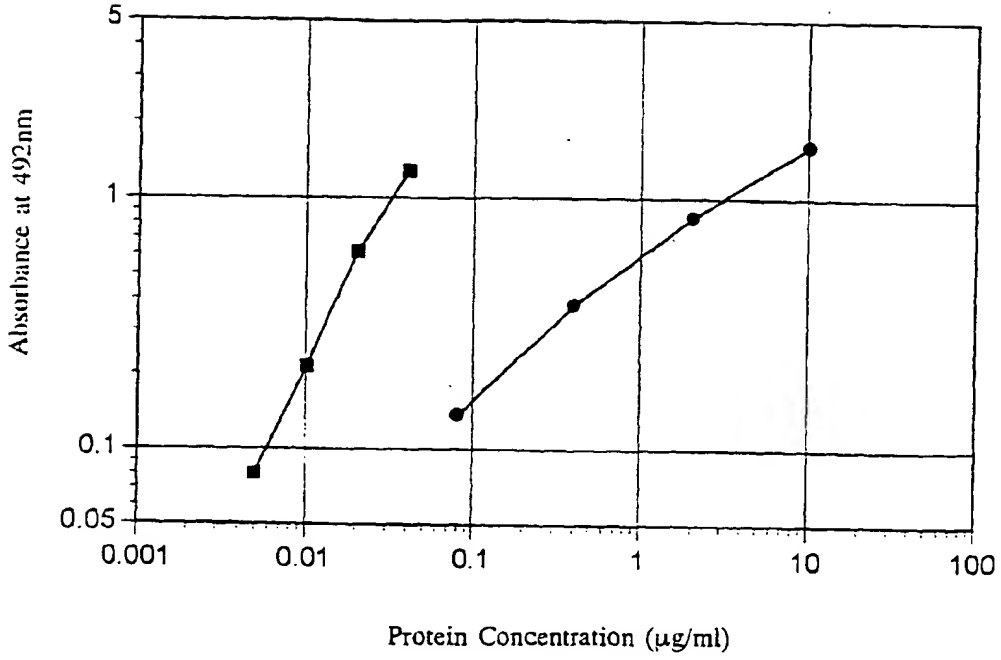
5
1
0.1
0.05
Absorbance at 492nm
0.001
0.01
0.1
1
10
100
Protein Concentration (µg/ml)

FIG.3